

# Pliny User Manual

## Table of Contents

1. Introduction
2. Getting Started
3. Using the Application
4. Understanding the Visualizations
5. Troubleshooting
6. FAQ
7. Glossary

## 1. Introduction

### 1.1 About Pliny

Pliny is a web-based application designed to visualize trends in Wikipedia metadata over time. Named after Pliny the Elder, a Roman author known for his encyclopedic work “Natural History,” this tool provides insights into what people around the world are reading, editing, and contributing to on Wikipedia.

### 1.2 Key Features

- Daily-updated visualizations of Wikipedia trends
- Historical data access dating back to September 2023
- Multiple visualization types showing different aspects of Wikipedia activity
- User-friendly interface for exploring metadata patterns
- Date selection for historical analysis
- Vandalism tracking and visualization

### 1.3 Why Wikipedia Metadata Matters

Wikipedia is the fifth most visited website globally, serving as a primary source of information for millions of people daily. By analyzing metadata (information about page views, edits, and other activity), Pliny reveals patterns in global information consumption and knowledge creation that would otherwise remain hidden.

### 1.4 Understanding Wikipedia Vandalism

As an open platform, Wikipedia faces challenges with vandalism—deliberate attempts to add incorrect information, delete valid content, or otherwise compromise the integrity of articles. By tracking and exposing vandalism patterns, Pliny promotes transparency and helps users understand how collaborative knowledge production works in practice. Our vandalism visualizations highlight which topics face the most challenges and how quickly the Wikipedia community responds to maintain accuracy.

## 2. Getting Started

### 2.1 System Requirements

Pliny is a web-based application that works in any modern web browser. For optimal experience, we recommend: - Chrome 90+, Firefox 88+, or Edge 90+ - JavaScript enabled - Screen resolution of 1280×720 or higher - Stable internet connection

### 2.2 Accessing Pliny

To access Pliny, visit [pliny.wiki](https://pliny.wiki) in your web browser. The application is freely available with no login required.

### 2.3 Interface Overview

The Pliny interface consists of several key elements:

1. **Date Selector:** Allows you to choose which date's data to view
2. **Visualization Area:** The main content area displaying various trend visualizations

## 3. Using the Application

### 3.1 Selecting a Date

1. Locate the date selector on the page
2. Click on the calendar icon to open the date picker
3. Select any date from September 1, 2023, to the present
4. The visualizations will automatically update to show data from your selected date

Note: Dates in the future or before September 2023 are not available for selection as data collection began in September 2023.

### 3.2 Interacting with Visualizations

Most visualizations in Pliny are interactive. Common interactions include:

- **Hovering:** Displays detailed information about specific data points
- **Clicking:** Selects a specific trend for detailed view
- **Zooming:** Some charts allow zooming for detailed inspection
- **Panning:** After zooming, you can pan to see different parts of the visualization

## 4. Understanding the Visualizations

### 4.1 Big Numbers (Aggregate Statistics)

This visualization shows total aggregate statistics across all of Wikipedia.

**How to read it:** - Each number represents a different metric (total views, total edits, etc.) - Values are calculated for the selected date

**What it tells you:** These numbers provide an overall picture of Wikipedia activity on a given day, showing the scale of engagement across the platform.

#### 4.2 Top Viewed Articles

This visualization shows the most viewed Wikipedia articles for the selected day along with their performance over a one-week period.

**How to read it:** - The line chart shows view counts over a one-week period - Multiple axes may be used to compare articles with different scales of popularity - Hover over points to see exact view counts for specific days

**What it tells you:** This reveals which articles are attracting the most attention and how their popularity has changed over the past week.

#### 4.3 Top Edited Articles

This radar chart array displays the articles with the most edits on the selected day.

**How to read it:** - Each radar chart represents a different article - The distance from the center indicates the number of edits - The shapes formed by connecting the points provide a visual comparison of their relative rankings

**What it tells you:** Articles with high edit counts often indicate topics that are actively evolving or controversial.

#### 4.4 Top Unique Editors

This bar chart shows which articles have the most unique editors contributing to them.

**How to read it:** - Each bar represents a Wikipedia article - The height of the bar indicates the number of unique editors - Hover over bars for exact counts

**What it tells you:** A high number of unique editors often indicates topics with broad interest or controversial subjects where many people want to contribute their perspective.

#### 4.5 Most Vandalized Articles

This bar chart displays articles that have experienced the most vandalism events:

**How to read it:** - Each bar represents a Wikipedia article - The length of the bar indicates the number of vandalism events - Hover over bars for additional details about the vandalism events

**What it tells you:** Heavily vandalized articles often reveal controversial topics, subjects of ideological conflict, or high-profile current events. This visualization exposes a hidden aspect of Wikipedia’s functioning—the constant battle to maintain information integrity.

**Why it matters:** By highlighting which articles face the most vandalism attempts, Pliny brings transparency to the Wikipedia editing process and helps users better understand which topics might be subject to manipulation attempts. This information is crucial for critical consumption of information and acknowledges the complex social dynamics behind Wikipedia’s creation.

#### 4.6 Most Grown/Shrunk Articles

This listing shows articles that have gained or lost the most content in terms of bytes.

**How to read it:** - Articles are listed with their title and byte change value  
- Positive values indicate content growth - Negative values indicate content reduction

**What it tells you:** Significant growth may indicate major new developments or information being added to topics, while reductions might indicate cleanup, consolidation, or removal of deprecated information.

#### 4.7 Articles with Most Gained Views

This line chart shows articles that have increased the most in popularity over the last three days.

**How to read it:** - Each line represents a different article - The y-axis shows view counts - The x-axis shows the three-day period - Upward trends indicate increasing popularity

**What it tells you:** Articles with rapidly growing viewership often reflect emerging news stories, viral trends, or topics gaining public interest.

#### 4.8 Articles with Most Lost Views

This line chart shows articles that have decreased the most in popularity over the last three days.

**How to read it:** - Each line represents a different article - The y-axis shows view counts - The x-axis shows the three-day period - Downward trends indicate decreasing popularity

**What it tells you:** Articles with declining viewership often reflect fading news stories, passing trends, or topics losing public interest.

## 5. Troubleshooting

### 5.1 Visualizations Not Loading

If visualizations fail to load: 1. Check your internet connection 2. Try refreshing the page 3. Clear your browser cache 4. Try a different web browser 5. Ensure JavaScript is enabled

### 5.2 Date Selection Issues

If you cannot select a specific date: 1. Verify the date is within the available range (September 2023 to present) 2. Try selecting a nearby date first, then navigate to your target date 3. Refresh the page and try again

### 5.3 Visual Display Problems

If visualizations appear distorted: 1. Try resizing your browser window 2. Check your zoom level (100% is recommended) 3. Disable browser extensions that might interfere with visualizations

## 6. FAQ

### 6.1 General Questions

**Q: Is Pliny affiliated with Wikipedia or the Wikimedia Foundation?**

A: No, Pliny is an independent project that analyzes publicly available Wikipedia metadata.

**Q: Does Pliny cost anything to use?**

A: No, Pliny is completely free to use.

**Q: How often is the data updated?**

A: Data is updated daily, typically with a 48-hour delay from real-time Wikipedia activity.

**Q: Does Pliny work with non-English Wikipedia editions?**

A: Currently, Pliny only analyzes English Wikipedia data. Support for other languages is planned for future releases.

### 6.2 Data Questions

**Q: Where does the data come from?**

A: All data comes from publicly available Wikipedia APIs and data dumps.

**Q: How far back does the historical data go?**

A: Currently, data is available from September 1, 2023, onward.

**Q: Is the raw data available for download?**

A: Pliny currently doesn't offer raw data downloads, but the visualizations reflect the processed data.

### 6.3 Technical Questions

**Q: Is Pliny open source?**

A: Yes, the Pliny codebase is open source and available on GitHub.

**Q: Can I contribute to Pliny?**

A: Yes, contributions are welcome. Visit the GitHub repository for details on how to contribute.

**Q: What technologies does Pliny use?**

A: Pliny uses Python for data processing, Go for the backend, and React/TypeScript for the frontend. Data is stored in Google BigQuery.

## 7. Glossary

**BigQuery:** Google’s enterprise data warehouse used by Pliny to store and analyze Wikipedia metadata.

**Edit:** A change made to a Wikipedia page, which can range from minor corrections to major content additions.

**Metadata:** Data about Wikipedia activity, such as page views and edits, rather than the actual content of articles.

**Page view:** A single instance of a user accessing a Wikipedia page.

**Spike:** A sudden, significant increase in activity (usually page views) over a short period.

**Trend:** A pattern of change in Wikipedia metadata over time that reveals insights about user behavior.

**Vandalism:** Edits made to Wikipedia pages that deliberately introduce errors or inappropriate content. Common forms include adding false information, deleting valid content, inserting offensive material, or spam.

**Reversion:** The act of reverting an article to a previous version, often done to counter vandalism.

**Visualization:** A graphical representation of data designed to make patterns and insights easily understandable.

**Wikipedia API:** Application Programming Interfaces provided by Wikipedia that allow access to their data.

**Z-score:** A statistical measurement used to identify how many standard deviations a data point is from the mean, used by Pliny to identify unusual activity.